

MICROCO PY RESOLUTION TEST CHART NATIONAL BUREAU H STANDARDS 1963 4

TIC FILE COPY AD-A182 953

SECURITY, LASSIFICATION OF THIS PAGE

	_
	7
	- 4
	J
(0)	

REPORT D	OCUMENTATIO	N PAGE			Form Approved OMB No. 0704-0188							
1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b. RESTRICTIVE	MARKINGS									
2a. SECURITY CLASSIFICATION AUTHORITY		3 DISTRIBUTION Approved	AVAILABILITY OF	REPORT release								
2b. DECLASSIFICATION / DOWNGRADING SCHEDU	LE		ion unlimit		•							
4. PERFORMING ORGANIZATION REPORT NUMBE	R(S)	5. MONITORING	ORGANIZATION RI	EPORT NU	MBER(S)							
6a. NAME OF PERFORMING ORGANIZATION University of California, San Diego	6b OFFICE SYMBOL (If applicable)	7a. NAME OF MO	DNITORING ORGA	VIZATION								
6c. ADDRESS (City, State, and ZIP Code) Department of Pharmacology M-03 La Jolla, CA 92093	36	76 ADDRESS (Cit	y, State, and ZIP (ode)								
8a. NAME OF FUNDING/SPONSORING ORGANIZATION U.S. Army Medical Research & Development Command	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT DAMD17-83	F FUNDING NUMBERS PROJECT. TASK WORK UNIT									
8c. ADDRESS (City, State, and ZIP Code)												
Fort Dietrick Frederick, Maryland 21701-5012		PROGRAM ELEMENT NO. 52734Å	62734A875	NO. AI	ACCESSION NO. 454							
11 TITLE (Include Security Classification) THE PRIMARY STRUCTURE OF ACETYL OF ORGANOPHOSPHATE TOXICITY	CHOLINESTERASE	AND SELECTIV	E ANTIBODIE	S FOR 1	THE DETECTION							
12 PERSONAL AUTHOR(S) Taylor, Palmer W., Ph.D.												
13a. TYPE OF REPORT 13b. TIME CO	OVERED 30/84 ₁₀ 9/30/85	14 DATE OF REPO December 3	RT (Year, Month, , 1985	Day) 15.	PAGE COUNT 18							
16. SUPPLEMENTARY NOTATION												
17. COSATI CODES FIELD GROUP SUB-GROUP O6 01	18. SUBJECT TERMS (
19. ABSTRACT (Continue on reverse if necessary The proposed work has been esterase. During the project to the top the esterase and a clone encoding for the 11S form of the enzyme and a C-terminal for the purpose of detecting the finally, professional esteration of the molecule. Finally, professional esteration is a second esteration of the molecule.	period we have ocid sequencing a m of the enzyme region have bea he phosphorylate	determined the and the isola . Peptides c en synthesize ed enzyme and	e primary sation and secorresponding and antib	tructum quencim g to the odies a g func	re of the ng of a c-DNA ne active center are being raised tional regions							
20 DISTRIBUTION/AVAILABILITY OF ABSTRACT		21 ABSTRACT SE	CURITY CLASSIFIC	ATION C	in the second							
UNCLASSIFIED/UNLIMITED SAME AS R	PT. DTIC USERS				<u> </u>							
22a. NAME OF RESPONSIBLE INDIVIDUAL Virginia M. Miller		301 663-7		²² SGR	FICE SYMBOL D-RMS							

THE PRIMARY SEQUENCE OF ACETYLCHOLINESTERASE AND SELECTIVE ANTIBODIES FOR THE DETECTION OF ORGANOPHOSPHATE TOXICITY

ANNUAL REPORT PALMER TAYLOR, Ph.D.

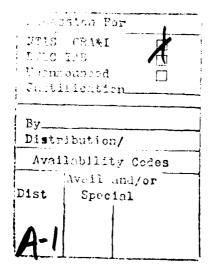
DECEMBER 3, 1985

Supported by

U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND Fort Detrick, Frederick, Maryland 21701-5012

Contract No. DAMD17-83-C-3202

University of California, San Diego La Jolla, California 92093





The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

SUMMARY

The proposed work has been directed to determination of the structure of acetylcholinesterase. During the project period we have determined the primary structure of the <u>Torpedo</u> enzyme through amino acid sequencing and the isolation and sequencing of a c-DNA clone encoding for the 11S form of the enzyme. Peptides corresponding to the active center of the enzyme and a C-terminal region have been synthesized and antibodies are being raised for the purpose of detecting the phosphorylated enzyme and delineating functional regions of the molecule.

FOREWORD

In conducting the research described in this report, the investigator(s) adhered to the "Guide for the Care and use of Laboratory Animals and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Animal Resources, National Research Council (DHEW Publication No. (NIH) 78-23, Revised 1978).

TABLE OF CONTENTS

Title Page	••••••	Page	1
Summary		Page	2
Foreword		Page	3
Table of Conte	nts	Page	4
Introduction	•••••	Page	5
Background	d	Page	5
Methods		Page	6
Results		Page	6
	o acid sequencing of <u>Torpedo</u> etylcholinesterase	Page	6
th	aration of antibodies directed to e active center of acetylcholin-		
	terase	Page	11
	arative sequencing of Torpedo and		
	her cholinesterases	Page	12
	ation of c-DNA clones encoding for etylcholinesterase	Page	13
st	ral aspects of acetylcholinesterase ructure deduced from amino acid and cleotide sequencing	Page	16
		Page	
	ist	Page	
Table I	Sequences of Torpedo californica	rage	10
Table 1	Acetylcholinesterase Peptides*	Page	8
Table II	Sequences of Action Site Regions	Page	12
Table III	N-Terminal Sequences of the Cholinesterase	Page	13
Figure 1	Active Center Sequences of the 11S and 5.6S Acetylcholinesterases	Page	9
Figure 2	High Pressure Liquid Chromatographic Profile of Acetylcholinesterase Prior to Following Endoglycosidase Treatment	Page	10
Figure 3	Acetylcholinesterase c-DNA Clones	Page	
Figure 4	Sequencing Strategy for Acetylcholinesterase	Page	
Figure 5	c-DNA and Interred Amino Acid Sequence of Clone 2-4	Page	

Introduction

During this period we have completed the primary structure of acetyl-cholinesterase through amino acid sequencing and c-DNA cloning and sequencing. This information should prove important to the many groups working on acetylcholinesterase structure.

Background

The control of the co

The high turnover number of acetylcholinesterase and availability of selective inhibitors have allowed extensive study of its distribution, catalytic mechanism, and functional role in cholinergic neurotransmission. The recent elucidation of the primary structure of the enzyme through protein chemistry and the isolation of its gene have added a molecular dimension to continuing studies on this protein, which controls the residence time of acetylcholine in the synapse. Acetylcholinesterase exhibits an extensive polymorphism of structure and, since the catalytic parameters of the individual enzyme species are largely invariant, the structural diversity appears critical only to the regulation of the cellular disposition of this molecule. Recent structural studies clearly show that acetylcholinesterase behaves as a secreted rather than an integral membrane The post-translational modifications provide an appropriate link to tether the enzyme to specific extracellular locations. Being an extracellular enzyme, modifications of structure critical to its disposition should occur prior to export to its site of residence. Thus, variations in structure responsible for cellular localization must either be encoded in the genome or be differentially affected by posttranslational events of biosynthesis.

Acetylcholinesterase Polymorphism

Since the initial finding of Massouliè and Reiger (1) that a native form of acetylcholinesterase contains an elongated tail unit linked to defined number of catalytic subunits, the control of individual species of acetylcholinesterase in relation to innervation, developmental processes and activity of excitable cells has received considerable attention (2). general classes of acetylcholinesterase species exist. The most unique is the elongated or dimensionally asymmetric species, which contains a filamentous tail unit disulfied-linked to tetrametic sets of catalytic subunits. The tail unit contains a collagen-like sequence distal to the catalytic Each strand of the triple helix is joined to a tetramer of subunits. catalytic subunits. Since each catalytic subunit is approximately 70,000 daltons, elongated species close to a molecular weight of one million are generated. In the case of Torpedo, but not Electrophorus, a second type of structural subunit has been identified as a non-collagenous, 100,000 daltoon peptide (3). It will be of interest if this structural entity also prevails in higher species. Treatment of the asymmetric form with collagenase markedly shortens the tail unit and a light tryptic digestion will remove the structural subunits without apparently altering catalytic parameters or the structure of the catalytic subunit (cf. 2). The asymmetric species appear to be fully assembled in the Golgi apparatus prior to export from the cell (4,5).

The second class are the globular forms, which show considerable structural variegation in subunit assembly (monomers to tetramers) and in hydrophobicity. The hydrophobic forms identified to date result from the cotranslational addition of glycophospholipid to the C-terminal carboxyl group of the nascent peptide chain (6,7). This modification resembles that seen in the variable surface glycoprotein of trypanosomes and the Thylantigen (8). It is quite possible that the nature of the glycophospholipid additions are not identical in the various tissues and may, in themselves, provide a basis for microscopic regional localization. Hence, the globular forms range from totally soluble species to species with particular hydrophobic glycophospholipids conjugated to the peptide chain.

Methods

The methods used for generation of the data described below have been documented in our manuscripts now published in the open literature and will only be described briefly.

- A. Determination of the primary structure of acetylcholinesterase Primary structure determinations relied on both tryptic and CNBr fragmentation. The peptides were initially size separated on Sephadex G-50 and then subjected to reverse phase HPLC on C-18 or C-4 columns. Sequencing initially involved dansyl-Edman and the Spinning cup, but after the first 6 months of the contract employed the gas phase method. Details may be found in (9).
- B. Antibody Generation and Assessment of Reactivity. Both monoclonal and polyclonal antibodies were made to the llS and 5.6S species of acetylcholinesterase. Similar methods were used for generation of antibodies to the individual peptides. Antibody reactivity and titers were determined by the enzyme-linked immunoassay (ELISA) method and by radioimmunoassay using $^{125}\text{I-acetylcholinesterase}$. Details may be found in references 3 and 10.
- C. Cloning and Sequencing of a cDNA-clone Encoding the 11S Species of Acetylcholinesterase. As described in reference 11, nucleotide probes to tandem sequences contained within a CNBr peptide were used to hybridize recombinants in a -gt 10 library. Positive clones were isolated, tested by hybridization and sequenced using M-13 sequencing vectors. Details are found in reference 11.

Results

estables account account account accounts. Assessed

1. Amino acid sequencing of Torpedo acetylcholinesterase

Our sequencing strategy is designed to fulfill several objectives:

- a. To obtain a sequence sufficient for the design of multiple nucleotide probes for c-DNA library sequencing.
- b. To employ the sequence to verify inferred amino acid sequence resulting for c-DNA sequencing and correlate the c-DNA sequences with the multiple acetylcholinesterase gene products.
- c. To elucidate differences in sequence between the molecular forms of acetylcholinesterase.

- d. To identify critical regions in the molecule: active center, chemically modified residues, N-terminal sequence, C-terminal sequence, glycosylation sites, cysteine-containing peptides and other sites of post-translational modification.
- e. To provide a peptide fractionation scheme by which other cholinesterases of lower abundance can be sequenced and homologous regions identified.

To date, we have sequenced about 80% of the tryptic peptides of the 11S enzyme and 30% of the tryptic peptides of the 5.6S enzyme. All of the CNBr peptides have been isolated from the 11S enzyme and 30% sequenced. A smaller number have been sequenced in the 5.6S enzyme. The sequences are summarized in Table 1. Several findings should be highlighted:

- a. Large tryptic peptides for the active center (24aa) were isolated and sequenced. Sequence was verified by chymotryptic digested and the position of this peptide in the whole enzyme could later be verified. The active center serine is serine 200. These peptides are identical in the 11S and 5.6S species (1) (fig. 1).
- b. The N-terminal peptides of the 11S and 5.6S enzymes were identified and sequenced through 42 and 30 residues, respectively. These peptides were also identical in the 11S and 5.6S species. These sequences were later verified by the c-DNA sequence and, more important, a leader peptide was demonstrated for the unprocessed acetylcholinesterase. Cleavage occurred C-terminal to an Ala, giving rise to the N-terminal Asp residue in the processed protein. A candidate C-terminal tryptic peptide ending in leucine was also identified in the 11S species. That this peptide was a C-terminal tryptic peptide was later verified by finding a stop codon at amino acid position 575, which followed the leucine code and thus ended the open reading frame on our cloned cDNA. An analogous C-terminal peptide has not been found for the 5.6S enzyme, and we believe a posttranslational modification occurred here, providing one of the points of structural departure of the two enzyme forms.
- c. The cysteine-containing peptides were identified by reduction and subsequent alkylation by $[^{14}C]$ iodoacetate. We obtained more cysteine peptides than would be predicted by the c-DNA sequence, but they arose simply from incomplete cleavages. All of these peptides can be placed in the inferred amino acid sequence on the basis of either their total sequence of their N-terminal residue identification and partial sequences. We have initiated fractionations of the unreduced enzyme with the essential aim of establishing the positions of the inter- and intrasubunit disulfied bridges. One of the eight cysteines appears to exist as a free sulfhydryl group and has been labeled with bimane in the 5.6S enzyme. Isolation and sequencing of the peptide reveals that the cysteine is at position 232.

Sites of glycosylation have been identified by lectin blotting of the individual peptides and by broad elution profiles that reflect microscopic heterogeneity within the peaks and their coalescence following endoglycosidase F treatment. An example is shown in fig. 2. Three of the four potential N-linked glycosylation sites have been located by peptide isolation (asparagine positions 56, 457, and 533), while it appears that position

TABLE I: Sequences of <u>Torpedo californica</u> Acetylcholinesterase Peptides*

11S Acetylcholinesterase		11S Ac	cetylcholinesterase (includes CNBr
146	ivgywa2fa-c	PCPUI	
177	vpvegcvfanef-nnci	bpd1	mwnpdre(p)
	fsivpvddgqfw(yst)k	bpd2	mnfvsnyypfgpgvylyflsieapd
1161	kpwsgvw-asnyp (carbohydrate	btl	mddnngiknrdglddivgdhnvicplm
	and CM cysteine)	bt2	m-wfg-p-pepgkpwngv-wasy-n
1161	kpwigvwfhnypl	bt3	mlntgnfkk(s)qillgvn(yk)(s)fgif(f)
IV33	dnhsellvntksgkvmgtrvpvlsshisafl		lyga(v)(g3)f
	givfaeqvgidv (N-terminal)	bt4	<pre>mhvwatfaktgnpnep-eg-(t)kwplifik-</pre>
IV 67	tvtifgesaggasvgmhilspgsr		(fq)-(e)
	(active site)	V187	vq-cwfwnqflp
IV 14	tgnpneptsqesk	VII17	
IV26	le-ea	V01	tgnpnephsqesk
IV62	fgbgtyly-pdtyr	V02	fidlntepmnk
IV63	ailqsg-vdcepa	V03	ailqsgspncpwasvsv
		V 04	galqwvhdniqffggdpmk
146	ivgywaa2fa-c		dedcly-niw-pgc
177	vpvegcvfanef-nnci		
V I 87	vqvcwfwnqflp	1161	
VII 17		IV26	hescael (C-terminal peptide)
V 02	fidlntepmnk	177	v-vegcvfanennci
V 04	galqwvhdniqffggdpmk	1101	d-nlvwpew-gvi-gy
167	iteahh	I 14 9	dlbbglncnl-nsaeelicl
1149	nlbbglncnl-nsaeelihicl	1159	lgvpda
11165	-(av)dedcly-niwspgca	1167	1-vphandlgldtvglqytdwmd
IV69	v-afalig	11188	f-ivpv-dgqfw
1167	12vphandlgld(5)v(g)lqytdwmddnngik		tgnpneptsqe
IV26	hescael (c-term)	1762	fgdgtyly
1161	kpw(i)gvw-as(n)ypl (carbohydrate,		aieag
1101	CM cysteine)	IV71	tvtifg-s
146	ivgywa2fa-(c)177	I4 6	ivgywa-fa
177	v(p)vegcvfane(f)(lp)nnci	167	iteah
	f(c)ivou(d)doofw(ve+k)	IV69	v-afali
	f(s)ivpv(d)dgqfw(ystk)	1103	4-01011
1163	dglddivgbhnvicplmhf		
1162	kpwphawdl g-p		
1164	lsvphandlgldlvt		
	dhnlvwpew-gvi(h)gyeig-l-p		
1168	lsvphandlgldtvglqytdwmd(ing)	5.6 A	cetylcholinesterase
	4e7 immunoreactive)		
173	v(aph)vegcvfane(yf)(lp)(np)nc(f)-	IV18	tgnpnep
	(hg)v(if)e	IX02	gpha-a
1149	nlbbglncnl-nsagglihicl (carbo)	V I 64	ail-epncpwatv-va
	-(av)(sd)edcly-niw(s)pgca	• • • •	dnhsqllvntksgkvmgt (N-terminal)
IV 57	ailqsgspncpwasvsv(aZg)r		tvtifgesaggasvgmhilspgsr (active
V 187	vq(v)cxfwnqflp		site)
161	kpw(1)gvw(f)(h)(n)y(p)1		
173			
	v-vegcvfanencfv(ip)g		
11164	b-dedcly-niw-pgc		
V65	ailqsgspncpwasvsv(azg)r		
V01	tgnpnephsqesk		
1154	lgvp-a		
IV 64	aieag		
IV71	tvtifg-s		

5.6S Acetylcholinesterase

```
dnhsqllvntksgkv-qt (N-terminal)
IV67
       tvtifa
1163
       dglddivgdhnvicplmhf
1164
       12 vphandlgl(dw)avt
V164
       a(i)lqsgsp(ns)cpwatv-va
      fgbgtyly(f)(f)n(h)r
tgnpnep(p)vzeq
IV62-64
ĬV 18
I X 0 2
       qpha-a
       aigag(a)(v)ae(pg)g-(v)-ppd
1g(v)(p)s(la)a--(dv)
IV 64
1154
I109
       (53)vd1(1)
      lgvpda---d--vp
1159
V 1 64
       a-lqsg(s)pncpw
1109
      (te)vd1(1)
1159
      1gvpda----d--vp
V164
       a-losg(s)pncpw
1167
       1gvphad-dq
       ailqsgsp(ds)cpwatv-va
V164
```

11S Acetylcholinesterase: CNBr Peptides

bt1 mddnngiknrdglddivgdhnvicp bt2 m-wfg-p-pepgkpwngv bt3 mlntgnfkk-qillgvn--fgif-lyga bt4 mhvwatfaktgnpnep-eg--kwplifik bd1 mwnpdre bd2 mnrvsnyypfgpgvylyflsieapd

*bt and bd peptides are CNBr peptides. Roman numerals denote Sephadex fractions of the initial fractionation and Arabic numbers denote the peak fraction from high-pressure liquid chromatography.

Sequence analyses of the active site tryptic peptides and chymotryptic peptide. The arrow indicates the [3H]isopropylphosphoryl-labeled senne. The consensus sequence is indicated in circles at the top. Residues in parentheses indicate placement only by amino acid composition. Residues indicated by acid composition. Residues indicated by capital letters are positions identified unambiguously, whereas sequences in lower case letters were tentatively identified. AChE, acetylcholinesterase.

115 ACHE TRYFTIC SPINNING CUP: THR- VAL THE-TLE-PHE-CLY-CLL R -ALA-CLY-CLY-ALA-SPE-VAL-CLY-MET-HIS-TLE-LEL SPE-PRO-CLY ISP- AF ... MANUAL : THE-VAL-THE-PHE-CLYCLY See Als Cly Cly ale See val Cly Met His Tie Les See Pro C.v. See Arch CHYHOTRYPTIC MANUAL : THR-VAL-THR-ILE-PHE-CLY(CIR Ser, Ala, CIV, CIV, Ala, Ser, Val, CIV, Net, Hi-) ILE-LEL-SER-PRO-CLY-SER-AR P4 P (1 de A peptide B 3.65 ACHE THYPTIC GAS PHASE: THE-VAL-THE-FILE-PHE-CLY-CLU- R -ALA-CLY-CLY-ALA-SET-VAL-CLY-HET-HIS-FLE-LEV-SET-PRC-CLY-(SET_AL) SOLID PHASE: THE - VAL -CLY - ILE - PHE - GIV - CIU-

Fig. 1 Active center sequences of the 11S and 5.6S acetylcholinesterases.

416, despite the presence of an Asn, X, Ser/Thr, is not glycosylated. Overall carbohydrate compositions suggest that we may have an O-linked site, but this remains to be established.

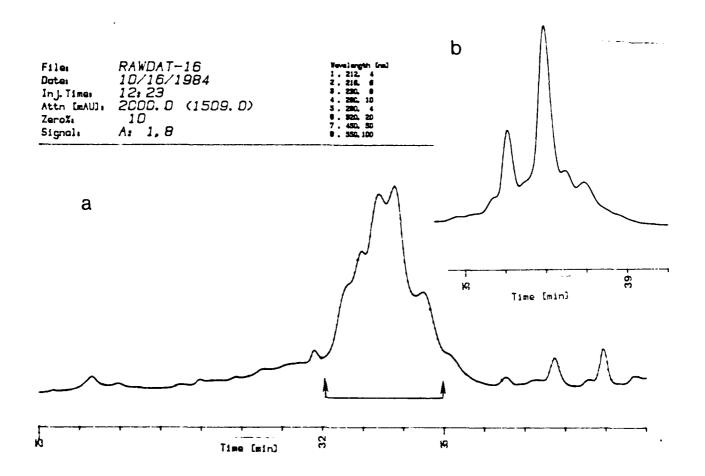


Figure 2: HPLC Profiles of 11S acetylcholinesterase peptides prior to (a) and following (b) endoglycosidase F treatment. Fractions 32-35 were isolated, lyophilized and treated with endoglycosidase F. The fractions were run on the same column (C-18 reverse phase), using an identical trifluoroacetic acid- $\rm H_2O$ -acetonitrile gradient. The elution profile with its altered elution positions and decreased complexity is shown in the inset.

CONTRACT STREET ASSESSED ASSESSED ASSESSED ASSESSED ASSESSED ASSESSED

d. Potential sites that serve as epitopes for the monoclonal antibodies raised by B.P. Doctor have also been identified. The two of particular interest are 4E-7 and AE-2. 4E-7 reacts selectively with the 5.6S enzyme (2) and has been found to react only with the glycosylated form

of the enzyme. Treatment with endoglycosidase F but not endoglycosidase H eliminates the antigenicity. 4E-7 reacts equally well with the native and denatured enzyme. A peptide extending between residues 358 and 386 shows the greatest reactivity with 4E-7 as determined by antibody blotting and competitive immunoprecipitation. We expect this peptide to be one of those unique to the 5.6S enzyme.

The other antibody of interest is AE-2, an antibody isolated by Fambrough and colleagues (3) which shows considerable species cross-reactivity. AE2 was found to react with a peptide found by B.P. Doctor in fetal calf serum acetylcholinesterase. This peptide has been found between positions 12 and 18 and considerable homology between species exists in a large portion of this peptide (cf: Table III). Several other antibodies are less well characterized. However, some, such as 4G-7 and 2C-9, show high titers and good immunoprecipitation capacity.

The rather brief description given here describes the bulk of the studies performed during the past 2 years. Extensive fractionation and sequencing were required to achieve this state of progress for an enzyme subunit size of 575 amino acids which exists in multiple enzyme forms. These endeavors have been very much facilitated by the instrumentation provided in the contract. Our basic sequencing strategy was to reduce and alkylate the protein with [14 C]iodoacetate and size-fractions were collected which were then subjected to reverse phase HPLC on $\rm C_4$ columns, using an acetonitrile-1% aqueous trifluoroacetic acid gradient. Peptides that fractionated poorly on $\rm C_{18}$ columns usually resolved well on $\rm C_4$ columns. Compositions and N-terminals were ascertained before subjecting the peptides to gas phase sequencing. Profiles of some of the many fractionations can be found in MacPhee-Quigley et al. (9).

2. Preparation of antibodies directed to the active venter for acetylcholinesterase

Having obtained the active center peptide sequence, we then synthesized a 25 mer peptide to generate antibodies to the active center of acetylcholinesterase. The peptide was synthesized by the Merrifield solid phase methods, using Dr. Russell Doolittle's facility, Department of Chemistry, University of California, San Diego. An N-terminal lysine was added to promote solubility, giving the sequence:

Lys-Thr-Val-Thr-Ile-Phe-Gly-Glu-Ser-Ala-Gly-Ala-Ser-Val-Gly-Met-Ile-Leu-Ser-Pro-Gly-Ser-Arg.

Antibodies are being generated in three ways: Monoclonals are being raised by B.P. Doctor and M.K. Gentry at Walter Reed. The fusions are now complete, positives colonies have been selected, and we should be screening for precise titers and selectivity next month. Polyclonal antibodies are also being raised in rabbits at San Diego, using two forms of immunogens: the peptide dispersed in liposomes and the peptide conjugated to hemocyanin. The monoclonal antibodies will have the potential of obtaining isolated antibodies directed to small peptidic domains, some of which show little species cross-reactivity. Other, owing to extensive homology, will exhibit considerable species cross-reactivity. The polyclonal antibodies can be expected to have the higher titers and will prove most useful for screening in vitro translation products and the development of highly sensitive assays

for the active center of acetylcholinesterase. Antibodies to synthetic peptides have the advantage of not showing cross-reactivity to contaminant proteins in biological preparations.

3. Comparative sequencing of Torpedo and other cholinesterases

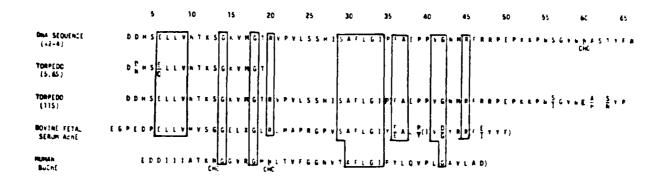
Our initial findings showing extensive homology of the active center peptides of Torpedo acetylcholinesterase and human butyrylcholinesterase (Table II) and the substantial homology in the N-terminal region of the two proteins prompted a further homology search in conjunction with Drs. Oksana Lockridge and Bert LaDu at the University of Michigan, and we see extensive similarity throughout the two molecules. Several peptides showing corresponding sequences can readily be found if our peptides (Table I) and their peptides are compared. The Michigan group also has about 80% of the peptides sequenced and with a total inferred sequence available in Torpedo, it should be possible for them to place the remaining peptides within the linear sequence.

TABLE II Sequences Of Active Site Regions

		_	••		•
Torpedo acetylcholinesterase	NH2 - THR VAL THR ILE	5 PHE GLY GLU	P 10 SER ALA GLY (15 GLY ALA SEF VAL GLY	20 <u>MET HIS 11E LEU SEF</u>
EEL ACETYLCHOLINESTERASE		era em	SER SER GLU (GLY ALA ALA GLY	
Human pseudocholinesterase	NH2 - SER VAL TIR LEL	PHE CLY GLU	SER ALA GLY	ALA <u>ALA SER VAL</u> SER	LEN HIS LEN LEN SEE
Equine pseudocholinesterase		PHE GLY GLU	SER ALA GLY	SER ALA ALA	
Equine Aliesterase		PHE GLY GLU	SER ALA GLY	ALA ALA SER	
BOVINE TRYPSINGEN	LYS ASP SER CYS	GLN GLY ASP	SER GLY GLY	PRO VAL VAL CYS SER	GLY LYS
PORCINE TRYPSIN	LYS ASP SER CYS	GLN GLY ASP	SER GLY GLY	PRO VAL VAL CYS ASN	GEA CEN
S. GRICEUS TRYPSIN	VAL ASP THE CYS	GLN <u>GLY</u> ASP	SER GLY GLY	PRO MET PHE ARG LYS	ASP ASA
F. COLL ALKALINE PHOSPHATASE	LYS PRO ASP TYP	R VAL THR ASP	SER ALA ALA S	SER ALA THR ALA TRP	SER THR

Human butyrylcholinesterase and <u>Torpedo</u> acetylcholinesterase can be expected to diverge on a phylogenetic basis and the basis of distinct enzymatic properties (i.e., the butyrylcholinesterase will accommodate substrates with large acyl groups, it does not show substrate inhibition and it is preferentially inhibited by different alkylphosphates). Therefore, one might expect that other mammalian acetylcholinesterases will possess structures showing structural divergence between these two limiting cases. In this regard, the fetal bovine serum acetylcholinesterase has proven useful. The trend in sequence divergence that we might expect can be seen in examining the N-terminal region of four cholinesterases (Table III). A more complete analysis of this mature should prove very useful in identifying various functional and antigenically cross-reactive regions.

Table III N-Terminal Sequences Of The Cholinesterases



4. <u>Isolation of c-DNA clones encoding for acetylcholinesterase</u>

Although this portion of the work was initiated and sustained with the support of the National Institutes of Health, the protein chemistry and molecular biological approaches are integrally linked, and it would have been impossible to proceed as rapidly without having both approaches in the Our library screening employed strategems that relies on same laboratory. hybridization with tandem but not overlapping probes, since we initially found that screening with a single probe yielded a very high incidence of false positives. When sequenced, the false positives were found to be repeating sequences of 500 bp with rather good base matches (14 of the 17 bases in the mixed probe). The tandem probes eliminated this artifact and were preferable to using probes coding for separate peptides. The latter approach will miss short length sequences. The tandem probe approach usually requires that more amino acid sequence be known, since rather long peptidic stretches are usually required to minimize code redundancy in the Positives to both tandem probes were then screened to a probe encoding for the N-terminal region. This reduced the number of positives and enhanced the likelihood of obtaining full-length inserts. By this approach we have now obtained 7 inserts which clearly encode for acetylcholinesterase and 13 more candidates. Their lengths and locations of Eco Rl sites are detailed in fig. 3. Only lambda 2-4 (AchE-1) has been fully The sequencing strategy (Fig. 4) and sequence (Fig. 5) are shown.

Acetylcholinesterase c-DNA Clones

(Eco Rl and Hind III Indicate Restriction Sites)

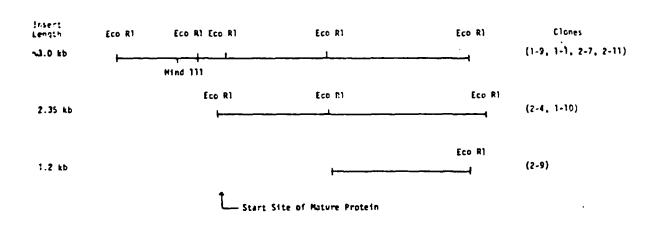


Figure 3: Characterization of several clones encoding for acetylcholinesterase. The length of the clones, Eco Rl and Hind III sites are shown.

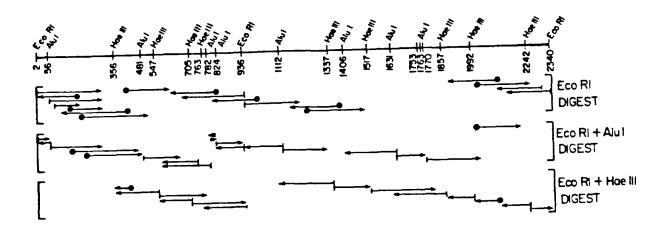


Figure 4: Sequencing strategy for acetylcholinesterase clone AchE-1 (γ 2-4). Critical restriction sites (Eco R1, Hae III, Alu 1) and sequencing primer sites are shown.

¥ 5	4 5 E:	Va1		D i	. .	\$2 k	S AGC	9 P P P P P P P P P P P P P P P P P P P	C.A.G 61.n	5 8 5	1726 : CAA	1788 BCAGBTCGCC	1838 C78CC788CC	1928 CETECTETAN	1998 CBGCCACCCA	2068 TCTCCACC10	2138 CACCTATTGA	2208 BTTCCCAACT	7278 AGCCAA1A10	
Ale P	1186 BAC AAC Asp Asp	A 80 0	- 2 Z	108 A 0.1		CAT TAC HIS TYT	6.88 8.18 8.18			CGC CAG	_ U =						Z C			
P 4	2 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	0 ±	1241 AAT 800 Aan 81 y	9 6 5 5	1396 CTG 678 Leu Val	E 5	C.A.G G1n G	2 4 E		6 6 6 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	AAG AAC Lys Asn	1778 BCTCTBCTCT	1848 CCTBBABACC	1918 CCCTCCCATC	1988 ACCAATCTT	2038 TATTACCCAC	2120 1CTCTCC181	2198 TCATCCCAGA	7248 1C18881648 2338	j S
S S		PASS ROC C	₽.	85 5	1 5	£ -	1301 10A C Ser G	F 3	35	1666 GCA GAA Ala Glu	100 t	XCTC.	100	7,70	4	HTR	:1016	:ATC	1186	
5 -	Tr B	287	₽ 0	1336 170 C	CTB CCT	A 004	100 THE SE	5 6 1 0		9 2 2	CAC T									
8 2	12 de 1	¥ -	AAB 111 80C Lys Phe 81y	1 91	800 C	F100	CCC C	ATT GAC 11# Asp	50.4	# c	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	C78 784	1838 CTAGTT	14 170CC	1. 18CG	2002	2 4	2188 CCCC01	2258 TACAAA 2328	Š E
¥ 3	ACA GAC Thr Asp	1231 ATC GTA GGG BAC 11e Val Gly App	1276 RCC A Thr L	1336 CTG BTG 1GG CCG Leu Val Trp Pro	11C	200 P	\$ 5 2 5	1346 111 A	10 1	ATT BAT 11e Asp	1711 ATG ATG Net Net	GA0 C	:AgTc	1 TOB GAGACTOCCT	1978 AATCTBCGGC	2048 ACTACCCCC	2118 CCCCTAAATC	2188 CCCTCCCC01	ccto	4 1 1
ATB TCA BBG BTC AND CTA ABC BTT CCC CAC Het Ber Biy Val Lys Lev Ber Val Pro His	2 2	중취	1276 TAC ACC Tyr Thr	4 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1381 010 1	C10 P	₽ C B		1606 CAG ATO 19C 61G Gin Met Cys Val	1651 PCC A	5 F	5 0	1998 1808 1808 1818 1838 166168ANGCA GAGACGGAG TCCGATATIG ACCAGACACC CAUTCTAGIT						2248 2258 TBRCCCTTCA CCTGTACAAA 2318 2328	
5	CTA CAG TAC Leu Gin Tyr	1216 AGA DAT BOA TTB DAC BAC ATG AND GIY LAU AND AND	¥ .	5 <u>t</u>	1381 77C 07C Phe Val	900 C	6 H C	CAG A	E S	1651 GAG ACC Glu Thr	AGT TCC TAC Ser Ser Tyr	1756 161 OCT Cye Ale	= 5 \$6 \$6	1848 CCTCGCCTTC	1968 OTCACTCGTA	2038 CTTCCCCACC	2108 CCATTCCAAT	7178 CCCAAG. ACT	2007 J	ž E
B >	₹ 3	200	Š.	1321 000 7	GAG 7	\$ n	MAC CCA	69 c	CGA GTT Prg Val	ACA 0	101	1756 3 AGC 1G1 5er Cya) U							
5	1136 136	1 40 I	GTT AAC	1321 CBA OCC TCA Arg Ala Ser	£ .	1426 046 6	600 4 7	AAB B	10 CTC 13	200	1696 166	200	1818 A11G	1888 1888	1 + 58	2028 TCCT	2048 3048	2168 ITCTC		
10 T B	1156 87C ACB Val Thr	ان اع ا	4 1 6	5 ±	GAG ATT Glu 110	1426 GAG GAG Glu Glu	غ <u>در</u>	1531 ACC A Thr L	CGA C	AAC GCC Aan Ale	1696 CGG TGG Arg Trp	CAC BAG Hrs 0to	BATA	וכפבי	1958 CTCTCGCACC	2028 1C1CC11CC1	2098 CCGTCAGGAA	2168 ACTETGTETE		אַ בַבְּבָּל
n i	PI V	6 6 5 4	1261 CAC 1111 His Phe	1306 CTB TAC TTC TTC AAC CAC Leu Tyr Phe Phe Aen His	1366 191 G	9C3 G	AMG ACT	ACT 13	C. 8.0 C	1636 CTC A	A s t	8 0	, 100	1978 1888 CCCCCCCA CCCCGCAAA					֖֭֓֞֞֞֜֜֞֞֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֡֓֜֜֓֓֓֡֓֡֓֡֓֡֓֡	Ē
U G	A 6			F. ŧ	1366 88C 1AT 61y Tyr	And A	1471 BCA A	n i	CAC C	1636 CTC CTC Leu Leu	Ė	. O. F	1808 CAGAG	1871 3000	7001	2016 MCTC	2086 30.017	21.5 300.1	222B 301C1 2298	
5 2	0 3	2 8 2	5 5	1306 11C 1	CAC HE & G	<u> </u>	2.5	CTC TTC Leu Phe	376 37.0 T	אשם ראיי ר	GAG TTT CAT Glu Phe His	TAC E	KBAGC	מככנ	1948 CCGACCTCGT	2018 AGCCCAACTC	2088 TCTCCCAATA	2138 CCCCCCC1C	77071	
00.7	1141 808 T	1201 F BOT ATA AND AND 1	CCC 118 A18 Fro Leu Met		0 +	= 5	٠ غ ي	הנד מ Pro L	1376 AAA 61C Lys Vel	55 t	1681 ACG 0	1741 GAC CAC TAC AGC / Asp His Tyr Ser /	6 2 9						E & B	2
5 5	= = =	본	1246 16C C Cye F	5	5.	E 2	E .	1516 156 C 179 P	Ato A	כלכ כככ Leu Pro	1681 AAG ACG Lys Thr	P C C	1798 3AAGCA	1868 SAGCCC	19) 1660	2008 BCATC	20) KCTC!	2148 :1ACCC	7218 MAAAGA 2288) 8
ATC TOT COO BAN DAC TIC TIP Ser Are Blu Asp Phe	1141 DAC TTA BOD TTO BAC AND LEU GLY LEU AND	P 10 0	1246 ATA 16C 11e Cye	401 174 C	419 1331 09C 07C 61y val	GAG CTG AAC 1	455 TGG GCG ACA Trp Ala Thr P	2. 4. 2. E. 4. 2.	CCC Pro H	304 1621 77C C	527 100 A		.es.	1868 CTCTGAGCCC	1938 CCCCAGGCTG	2008 CTCTCBCATC	2078 CCGCACTCTC	2148 COTTCTACCC	7218 2228 2238 CCCCAAAAGA CTICTIGGICI CCTICCTAAG	
					_					_						_				
5 ki	106 CCA BTC Pro Vel	£ 3	10 p	נו ב	181 Cys	40.4	1 PC		90 A	ATC -	646 637 617	7. A	110			= 3	110			5 § .
BAT CAC TET	6 r	5 5	211 670 Vel	£ £	4 4	87C	0 to v	# 5 P P P P P P P P P P P P P P P P P P	A 5	\$ £	9 6	7 4	£ 7.			200	100	ATB 119	1021 1021 BAC GAG	300
BAT CAC	0 TC	100	93A 91 y	5 E	289	Ş	2 5	5.5	9 4 1	8 5 4 0 5 4	₽ 5 8	5 3	910	191	C Y		110			Lys Amp
2 9	ACA AGA BTC Thr Arg Vel	151 GAG G14	108 501	2 E	98	A P	121	101	9 0	S I	8 8	6.9 0.50 0.50 0.50	Į,	§		7.2	BU 1			֓֞֞֞֞֜֞֞֞֞֞֜֞֞֜֞֜֜֞֜֞֜֜֞֜֜֞֜֓֓֓֓֓֓֓֓֓֓
Cys Gin Ale Ang		9 e	1 g 7	4 5 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	5 1	20 T	€ +	91C	326 CAG GI n	070 147	\$ 2 21 S	§ 8	910	2.0 CTC	3	ANG ANG Lys Lys	110	§		000 000
8 €	418 804 Het 81x	E		16 ×	81 u	1 8 8 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	5	55	A 00.	108	35 %	Ser	188	. §	A C	₹ <u>₹</u>	901 A10		5 5	· §
Ω <u>6</u>	E E	7 °	: § :	5 Y	§ §	5 °	ž Š	464 010 1-0	9 60 61 v	C.A.G	A =	900	736	. §	۶	\$ 2	1 0		00 E	# 4 A
	5 2	£ =	\$ 5	6 <u>0</u>	0 8 4 0 8 4	804 5	800 61 y	810 1010	S I	S71 CTG	Ž Ž	5 C	16C	S	4 5	P 60	9		E 10	3 2
2.2 2.3	₹ 5	136 808 61 y	BAG CCC AAB	241 10C CAB CAB TAC Cye Gin Bin Tyr	301 AAT CCG AAC Aan Pro Aan	CC8 A88 Pro Arg	2 2 3	2 5	CTC CAC	60 € • 1 •	ACB BTG ACC Thr Val Thr	676 100 50r	AAT TGC	781 0A8 CT8 00A /	3	181 CT8 AGG Cys Leu Arg	CCC 111 GAC	2 50	TTA C18	2 5
5 ÷	4 M	53	3 5	241 70C Cye	\$ \$	į į	\$ 7	3 5	E 0 4	£ 1		문로	9		20	181 Cys	ננ נ	ב ב	F 2	500
ē >	\$ 50 ¥	E E	CC7	3 5	E c	\$5.2 \$5.2	E &	1	5.3	80 g	\$ \$ £	E .	Ď,	910	3	S E	9 - 1	2 2	£ 5	
د ء	₹ 5	136 8CT TTC CT8 808 ATT CCC Ala Phe Leu Gly Ile Pro	± § ₹	1 5	2 1	6 ×	967 9 1 y	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	E	8 5	55	¥ ±	721 : 80C TC8 CCC / 81 y Ser Pro /	909	ē	14 E	gtc	. §	991 ACT CAG ATC	1051 1051 1064 BCB CCB 007 TTC
1) I	Ž Ž	9 5	\$ \$	S S	286 GAG ATB 188 (Blu Met Trp /	8 5	900 91 y	9 <u>-</u>	80 Y	33.6 GAC Pep	30	676 ATB CAC ATT CTC TCC CCB 600 Met His Ils Leu Ser Pro 61y	ă,	9	2 2	5 3	AAT 840	946 908 BAA TTC TTC CCA ACC	A 15	ָּצָּ בַּ
E 3	¥ 5	12 ·	E &	1 ×	ž į	346 ATT THE GTM CCT IIe Trp Val Pro	1 8 A	E 3	E	D 2	980 31 Y	188	8	8	ž	200	90	. F.	£ §	311 1034 171 170 178 178 170 170 170 170 170 170 170 170 170 170
AAT TCC 818 CTT CTG CAC TTG 81C 81C CTG Asn Ber Val Leu Leu His Leu Val Val Leu	BNS CTC CTS STC AND ACC ANS TCG SGA AAA STC SIL LEU LEU VAI Ann Thr Lys Ser SIY Lys Val	CAC ATC ABC (His He Ser A	8 5	23.8 17.	BBY Ser	¥ £	341 ATC TAC GBA BDC BBT TTC TAC 110 Tyr Bly Bly Phe Tyr	113 434 434 431 432 431 600 146 400 Am 61y Lys Tyr Leu Ala Tyr Thr	496 878 GGC GC1 111 GGT 111 Val 81y Ala Phe Gly Phe	936 GTC CTG GAC CAB ACB ATG BCA CTG BIY Lev Lev Aup GIn Arg Met Ala Lev	167 601 616 CAS 17C 17C GBC BBG BAC CCC AAG Bln Phe Phe Gly Gly Asp Pro Lys	6401 681C BCC TCT BTC BBC BLY Ala Ser Val GLY 201	706 GCC ATC CTT CAB ADC Ata 11e ten Blo Ser	764 COC AGO AOG BCB BTC	į	BAC BAA BAG CTC ATC CAC Asp Blu Blu Leu Ite His	871 878 848 708	ב ב	ָבָּ בַּי	22
5 × 5	4 E 3	¥ \$	21	8	្ត រួ	22.2	2	§ 5	2	3 4 6	10 T T T T T T T T T T T T T T T T T T T	2 3	100	## 88C C	À 15	3 3	370	CCC 01C ATC	MAC 110	
Ų.	۲ کا ا	5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	3 ¥ £	9C4	271 271 88A 111 1 61y Phe 5	178 18C S	010 198 / Val Trp I	387	2 2	0 4 6 4 6	ñ.	88C 8CC 7	88.5	. 5	2	Ser	A11 BAC 0	. א א	476 996	- F9 > <u>F</u> 9
F F	듔삒	2 3 1	-22	# Z	<u> </u>	<u> </u>		■ → ·		¥ 2	□ #	Z Z1	5.4 -	3	=	* -	, <u>, , , , , , , , , , , , , , , , , , </u>	بر		2

Figure 5: c-DNA and inferred amino acid sequence of clone 2-4.

The insert begins with a 16 amino acid leader peptide, extends through the 575 amino acids of the processed protein, and contains another 570 bases in It does not contain a poly A termination or a a 3' noncoding region. canonical poly A initiation signal, which suggests that this 3' region is not complete (4). Clone 2-4 was sequenced in M-13 by the dideoxy method. Protein sequence, again, provided confirmation that the selected open reading frame was correct and did not divigate due to a skipped base. addition, the protein sequence enabled us to establish that the clone likely encoded for the 11S species. We are sequencing the other clones based on initial findings, are optimistic that we have found a clone for another acetycholinesterase species. There is one caveat: The fact that all clones end in EcoR-1 sites suggests incomplete methylation in the library prepara-Clones 1-1, 1-9 and 1-10 are probably identical but reflect another gene of acetylcholinesterase. They are being sequenced. Clone 1-9, which is nearly 2.9 kb in length, is our candidate for obtaining complete 5' and 3' noncoding regions. Clone 2-9 is probably a shortened version of 1-9, terminating at the EcoR-1 site. In short, valuable information will continue to accrue as we compare c-DNA inferred sequences with actual protein sequences. Accordingly, the combination of molecular biology and protein chemistry should enable us to identify all of the structural polymorphisms in Torpedo acetylcholinesterase.

5. General aspects of acetylcholinesterase structure deduced from amino acid and nucleotide sequencing

All of the above data enable us to arrive at the following conclusions:

- a. Acetylcholinesterase contains a hydrophobic leader sequence (residues -13-0) but contains no other hydrophobic domains which are candidates for membrane-spanning regions. This it is likely to be an exported protein and its membrane attachment site(s) arises as a consequence of posttranslational modifications.
- b. The active center serine is at residue 200. The N-terminal location contracts with the serine proteases of similar size that function in the clotting cascade (i.e., factor IX and prothrombin).
- c. No significant global or local homology is found with the acetyl-choline receptor.
- d. Although acetylcholinesterase is closely homologous to human butyrylcholinesterase, no significant global homology and very limited local homology are found with other serine proteases: the largest local homology is seen with liver aliesterase and the carboxylesterases.
- e. Substantial homology is found between acetylcholinesterase and thyroglobulin in their C-terminal regions (acetylcholinesterase residues 1-575; thyroglobulin residues 2168-2750). Six of the eight cysteines are conserved, suggesting a similar folding pattern for the two macromolecules. The region between 160 and 190 shows greater than 60% identity.

REFERENCES

- 1. Massouliè, J., and Reiger, F. (1968) Eur. J. Biochem. 11, 441-455.
- 2. Massouliè, J., and Bon, S. (1982) Ann. Rev. Neurosciences 5, 57-106.
- 3. Lee, S.L., Heinemann, S., and Taylor, P. (1982) J. Biol. Chem. <u>257</u>, 12302-12309.
- 4. Rotundo, R.R. (1984) Proc. Natl. Acad. Sci. USA 81, 479-483.
- Younkin, S.G., Rosenstein, C., Collins, P.L. and Rosenberry, T. (1982)
 J. Biol. Chem. 257, 13630-13637.
- 6. Futerman, A.M., Fionini, R.M., Roth, E., Low, M.G. and Silman, I. (1985) Biochem J. 226, 369-377.
- 7. Rosenberry, T. (1985) Biochem. Biophys. Res. Commun. 133, 621-627.
- 8. Low, M.G., Ferguson, M.A.J., Futerman, A.M., and Silman, I. (1986) TIBS 11, 212-215.
- 9. MacPhee-Quigley, K., Taylor, P. and Taylor, S.S. (1985) J. Biol. Chem. 260, 12185-12189.
- 10. Doctor, B.P., Gentry, M.K., Camp, S., Taylor, S.S. and Taylor, P. (1983) Proc. Natl. Acad. Sci. (USA) 18, 5767-5771.
- 11. Schumacher, M., Camp, S., Maulet, Y., Newton, M., MacPhee-Quigley, K., Taylor, S.S., Friedmann, T. and Taylor, P. (1985) Nature, 319.
- 12. Fambrough, D.H., Engel, A.G. and Rosenberry, T. (1982) Proc. Natl. Acad. Sci. USA 79, 1078-1082.

DISTRIBUTION LIST

4 copies Commander

US Army Medical Research and Development Command

ATTN: SGRD-RMS

Fort Detrick, Frederick, Maryland 21701-5012

5 copies Commander

US Army Medical Research and Development Command

ATTN: SGRD-PLE

Fort Detrick, Frederick, Maryland 21701-5012

12 copies Defense Technical Information Center (DRIC)

ATTN: DTIC-DDAC Cameron Station

Alexandria, VA 22304-6145

1 copy Dean

School of Medicine

Uniformed Services University of the

Health Sciences 4301 Jones Bridge Road Bethesda, MD 20814-4799

1 copy Commandant

Academy of Health Sciences, US Army

ATTN: AHS-CDM

Fort Sam Houston, TX 78234-6100